Name: Sourabh Mehta
Phone: +91- 9110904883
Email ID: sourabhmehta06@gmail.com

SUMMARY

PG Diploma in Data Science and Engineering has an interest in programming and data-driven solutions to increase efficiency, accuracy, and utility of internal processing.

- GitHub: https://github.com/SourabhMehta-sm
- Kaggle: https://www.kaggle.com/sm12120552
- LinkedIn: https://www.linkedin.com/in/sourabh-mehta-667959174/

KEY SKILLS

- Python, MySQL, Tableau, Statistics
- EDA, Data cleaning, Missing value treatment, Outlier treatment, and Modeling
- Machine Learning:
  - Supervised Learning: Decision trees, Random Forest, Linear Regression, Logistic Regression, KNN
  - Unsupervised Learning: K-Means Clustering, Principal Component Analysis.
- Web scrapping

ACADEMIC PROJECTS

- **IBM HR Analytics Employee Attrition & Performance**
  This dataset consists of 1470 employees details which consist of 35 features out of which 18 features were continuous and 17 features were categorical. The cost function/target variable is attrition. Drained some conclusion from an analysis of data that some features were affecting the attrition a lot such as monthly income, over time, marital status, etc. Attrition contains binary data, so build a classification model, used logistic regression as baseline model and final model as KNN.

  Key skills:
  Performed data pre-processing, exploratory data analysis, splitting data in 80:20 train test, KNN

- **Indian lever patients analysis**
  This dataset consists of 582 patients/observations with 11 features providing details about patients. There are 9 continuous features and 2 categorical features including the target/cost function which is whether the patient is suffering from liver disease or not. Analysis related to the target variable was done to know the relationship between the dependent variable and the independent variable. As it is a classification so build a KNN using 6 number of nearest neighbours model to classify.

  Key skills:
  Performed data pre-processing, exploratory data analysis, feature selection using recursive feature elimination, build a KNN model.

- **Rossmann Store Sales**
  Data related to Rossmann stores and their customers which contains 10lakhs entries and 18 features. So the main aim was to predict sales of every Rossmann stores across the country, applied certain machine learning algorithms(Regressor) to build a predictive model to predict sales, found Decision Tree Regressor as best.

  Key skills:
  Data preprocessing, ML.

Name: Sourabh Mehta
Phone: +91- 9110904883
Email ID: sourabhmehta06@gmail.com

EDUCATION

Course	Institution	Year	Remarks
Data Science and Engineering	Great Lakes Institute of Management	2019	Completed
B. Tech	Guru Nanak Dev Engineering College, Ludhiana	2019	61.1%
12th Std	NIOS	2015	74%
10th Std	St.Karens High School	2012	80%

OTHER ACHIEVEMENTS

- Industrial Training at Indian Railways in Signal & Telecommunication department.
- Build a Ham Radio.